

# ON VERIFICATION OF UPPER-AIR WINDS BY VERTICAL SHEAR AND EXTREMES

OSKAR M. ESSENWANGER\*

Research Laboratory, Army Ballistic Missile Agency, Huntsville, Ala.

ROBERT E. BRADFORD

National Weather Records Center, U.S. Weather Bureau, Asheville, N.C.

and

WILLIAM W. VAUGHAN\*\*

Marshall Space Flight Center, National Aeronautics and Space Administration, Huntsville, Ala.

[Manuscript received November 23, 1960; revised February 17, 1961]

## ABSTRACT

The existence of undetected errors in recorded wind observations may have a biasing influence on a statistical study. In the progress of some studies it has been found necessary to reexamine the data being used. A series of upper-air winds has been checked by using available listings of vertical shear and extreme winds. The developed procedure permits correction for major errors and tolerates the minor (random) errors.

The test of data by maximum wind profiles uses the highest and second highest scalar wind speed for each station and checks the data by profile scan. The test of data by vertical wind shear uses a critical value, theoretically derived, exceedance of which marks the data as suspicious. A detailed check of the wind observation verifies this suspicious value or it is corrected. In this program 3.5 percent of the observations<sup>1</sup> proved suspicious and 85 percent thereof, that is, 2.9 percent of the observations, required correction. Thus the critical value is highly efficient.

The errors were traced and split into clerical errors (1.1 percent), instrumental errors (1.3 percent), and computational errors (0.5 percent), which are quite within reasonable limits.

## 1. INTRODUCTION

For use in missile design and performance studies by Army Ballistic Missile Agency, basic upper-air wind observations were obtained for locations in the Pacific Ocean, North America, and Europe. The stations are listed in table 1. Preliminary analysis of the data revealed that there were occurrences of apparent errors in the observations as presented on punched cards. It was decided that these data should be checked.

Although it was considered desirable to check and verify all the upper-air data, this was not possible because of the cost and the time required to review the mass of observations. Instead it was decided to establish a checking program which would permit a review eliminating major errors, yet tolerating minor (random) errors. It was considered to be sufficient to restrict the checking process to higher wind magnitudes and wind shears, where the possibility exists that the reported extreme wind velocity arises from the addition of wind data and error with the same sign.

This method permitted the correction of the major items

(maximum wind speed and wind shear) required for missile design studies and verification of the wind data at the same time.

The mass of punched cards was converted to magnetic tapes for use on high-speed electronic computers (IBM 704 and 709). Use of these computers permitted the rapid searching of the data and machine listings of all observations (plus the associated profile) which produced the higher wind speeds and wind shear for each altitude level. Also, it was possible to provide preliminary frequency distributions of the wind shear and speed data for use in further evaluation of suspicious data. The corrected observations were subsequently incorporated into the original data records and utilized on various statistical programs for use in missile design and employment studies.

## 2. THEORETICAL BACKGROUND

The problem existed that frequency distributions for wind shears and extreme values had been programmed, and tabulations similar to table 2 (described later) had been made before the necessity for critical review and correction became apparent. Thus, the problem was not to establish a suitable statistical theory of fitting extreme value data, but rather to develop an economical tech-

\*Formerly associated with the National Weather Records Center, U.S. Weather Bureau, Asheville, N.C.

\*\*Formerly associated with the U.S. Army Ballistic Missile Agency, Huntsville, Ala.

<sup>1</sup> "Observation" used here refers to the entire ascent though sometimes only the value at a single level had to be corrected.

TABLE 1.—Station listing of verified wind data <sup>1</sup>

Job station No.	Station name	WBAN No.	Period of record	Hour <sup>2</sup> (GMT)	Type of equipment <sup>3</sup>
1	Cocoa (Patrick AFB).....	12867	1/51- 9/52	00, 12	4, 6, 8
			10/52-11/56	00, 12	8
	Cape Canaveral.....	12868	11/56-12/57	00, 12	8
2	Tokyo (Haneda AFB).....	43311	1/51- 4/52	00, 12	2, 6
			4/52-12/53	00, 12	8
			7/55-12/55	09, 21	8
			8/57-12/57	06, 18	8
	Nagoya.....	43312	1/54- 6/55	00, 12	8
			1/56- 5/57	00, 12	8
3	Fairbanks.....	26411	1/51-12/57	00, 12	2, 6
4	Tripoli (Wheeler AB).....	33123	1/51- 4/53	00, 12	8
			5/53-12/57	00, 12	8
5	Berlin.....	35140	1/51- 6/55	00, 12	8
6	Wiesbaden.....	35010	1/51- 3/53	00, 12	8
			4/53- 6/55	00, 12	8
7	Bitburg.....	34049	1/56-12/57	00, 12	8
8	Port Lyautey.....	13017	1/51-10/51	00, 12	6
			12/51-12/52	00, 12	6
			1/53-12/57	00, 12	8
9	Thule.....	17605	1/51- 3/54	00, 12	8
			5/55	00, 12	8
			1/56-12/57	00, 12	8
10	Lihue.....	22536	1/51-12/57	00, 12	8
11	Grand Bahama Isl.....	12712	1/51-12/57	00, 12	8
12	International Falls.....	14918	1/51- 9/54	00, 12	6
			10/54-12/57	00, 12	8
13	Guam (Anderson AFB).....	41414	1/51- 4/53	00, 12	8
			5/53-12/57	00, 12	8
14	San Juan.....	11631	1/51- 5/55	00, 12	6
			1/56-12/57	00, 12	6
15	Balboa, C.Z. (Albrook AFB).....	10701	1/51-11/53	00, 12	8
			11/53- 4/55	00, 12	8
			6/55- 7/55	00, 12	8
			9/55-12/57	00, 12	8
16	Johnston Isl. AFB.....	21601	1/51- 4/53	00, 12	6, 8
			5/53- 2/54	00, 12	8
			3/54-12/57	00, 12	8
17	Luzon (Clark AB).....	41207	1/51- 8/53	00, 12	6
			9/53-12/57	00, 12	6
18	Santa Maria.....	23236	1/51-10/54	00, 12	6
		23273	11/54-12/57	00, 12	6
19	Keflavik (Meeks Fld).....	16204	1/51- 8/51	00, 12	8
		16201	9/51-12/57	00, 12	8
20	Silver Hill (Washington, D.C.).....	93722	1/51-12/52	00, 12	6
			1/53-12/57	00, 12	6
21	Adak.....	25704	1/51- 5/53	00, 12	8
			6/53-12/57	00, 12	2
22	El Paso.....	23044	1/51- 4/54	00, 12	6
			5/54-12/57	00, 12	6
23	Barrow.....	27502	1/51-12/57	00, 12	6
24	Alert.....	18601	7/53- 6/58	00, 12	6

<sup>1</sup> The sequence of stations is not identical with that in tables 2 or 4.

<sup>2</sup> Prior to May 1957 observations were taken at 03 and 15 instead of 00 and 12. This is not listed here.

<sup>3</sup> Type of equipment: 2=Rabal, single theodolite

4=SCR-584 and SCR-545

6=SCR-658

8=GMD-1 and GMD-1A

nique with minimum amount of time and cost using these available tabulations for selection and checking procedures.

From the statistical point of view it must be pointed out that the problem of establishing profiles of maximum wind speed (magnitude of wind vector) or distributions of maximum wind shears should be approached otherwise than shown in this report by employing an extreme value statistic. However, frequency distributions of maximum wind speed and wind shear in the upper air are not completely known and application of techniques similar to Thom's [2, 3] method or Gumbel's [1] theory would have made necessary the careful evaluation of several such statistical systems. This time-consuming basic study could not have been completed within the limitation of the available time and funds.

The first part of the verification program, checking maximum wind profiles, was relatively simple. The listed and plotted profiles were scanned for suspicious values. Verification of those values considered suspicious was per-

formed by reference to the original data records. Details of the procedures employed are presented in section 3.

The selection of suspicious values for wind shear data appeared very complicated in the beginning. Distributions of wind component shear frequencies were available in the form of table 2 without the column marked "Essenwanger's sum". In the form shown in table 2, frequency distributions were given without regard to algebraic sign of the wind shear. The problem was to find a value which separated the acceptable values from suspicious values without reviewing too many observations or accepting a large amount of unreliable data. This value will be called the critical value  $\epsilon_c$ .

The column marked "99.865" was practically identical with the maximum shear value. There is no reason to expect all maximum shear values to be wrong. On the other hand, the value listed in the column marked "97.72" was assumed to be acceptable since regularly it was exceeded by 2 out of every 100 values. Thus we may contemplate the following ideas.

We start with the assumption that zonal and meridional wind components are normal (Gaussian) distributions, or approximately normal. Departures from normality will be introduced later in the discussion. The wind shear data must then also follow a normal distribution. The shear distribution, disregarding the sign, then is a folded distribution. This folding occurs at the zero wind shear value. The mean value of the frequency distribution of zonal or meridional shear values, not disregarding the sign, usually will not coincide with zero. The question is now: Which portion of  $\sigma$  (standard deviation) corresponds to the listed percentage frequencies, 50, 84.1, 97.72, etc.?

If the folding occurs outside the  $\pm 3\sigma$  value <sup>2</sup> in reference to the mean, then practically all shear values have the same sign. Then the shear distribution follows a normal distribution and the 50 percent value (median) virtually coincides with the mean. This statement holds for folding above  $\pm 1.5\sigma$  (fig. 1a) if we assume that the frequency of data above  $3\sigma$  is negligible. If the folding is within  $\pm 1.5\sigma$ , then the mean value of the unfolded normal distribution must be smaller than the 50 percent value (fig. 1b). Thus the listed 50 percent value permits evaluation of the magnitude of this mean value. We recognize (see also table 2) that for practical purposes this value is so close to the zero wind shear that we can continue our discussion about the folded distribution as if it were folded at the mean value zero. Then the 84.1 percent value corresponds to  $1.41\sigma$ , the 97.72 percent to  $2.28\sigma$ , and the 99.865 percent to  $3.20\sigma$ .

We build the ratio

$$\frac{99.865 \text{ percentage value}}{97.72 \text{ percentage value}} = \frac{3.20\sigma}{2.28\sigma} = 1.40$$

Thus, theoretically we should expect the factor 1.4. A

<sup>2</sup> Some statisticians may want to use the symbol  $s$ , the estimate of the population standard deviation  $\sigma$ . This has no influence upon the development in this paragraph insofar as later the 84.1 or 97.72 value may express the empirical replacement for  $\sigma$  or  $s$ .

TABLE 2.—Meridional wind shear ( $\text{sec}^{-1}$ ) distributions at Keflavik, Iceland, for January. Station altitude, 54 m. MSL; latitude,  $63^{\circ}57' \text{ N.}$ ; longitude,  $22^{\circ}37' \text{ W.}$ ; period of observations, January 1951–December 1957

Altitude km. (MSL)	Number of OBS	Cumulative Percentage Frequency								Essenwanger's sum	Max shear	Pct. freq.	Suspicious values
		.135	2.28	15.9	50.0	68.0	84.1	97.72	99.865				
SFC-0.5	372			.0026	.0090	.0131	.0202	.0369	.0808	.0621	.0808	.27	*
0.5-1.0	370			.0008	.0037	.0055	.0093	.0199	.0291	.0342	.0291	.27	
1.0-1.5	368				.0030	.0055	.0092	.0227	.0314	.0369	.0314	.27	
1.5-2.0	369			.0001	.0028	.0046	.0090	.0176	.0343	.0316	.0343	.27	*
2.0-2.5	372				.0028	.0042	.0077	.0153	.0306	.0280	.0306	.27	*
2.5-3.0	374				.0037	.0057	.0085	.0204	.0319	.0339	.0319	.27	
3.0-4.0	375			.0004	.0027	.0040	.0061	.0103	.0163	.0214	.0163	.27	
4.0-5.0	373			.0003	.0021	.0038	.0064	.0139	.0267	.0253	.0267	.27	*
5.0-6.0	371			.0004	.0027	.0049	.0080	.0164	.0274	.0294	.0274	.27	*
6.0-7.0	359				.0023	.0042	.0069	.0166	.0656	.0285	.0656	.28	*
7.0-8.0	351			.0002	.0025	.0042	.0074	.0157	.0333	.0281	.0333	.28	*
8.0-9.0	335				.0028	.0043	.0074	.0159	.0480	.0283	.0480	.30	*
9.0-10.0	303				.0023	.0038	.0074	.0172	.0345	.0296	.0345	.33	*
10.0-11.0	278			.0004	.0028	.0049	.0088	.0243	.0390	.0381	.0390	.36	*
11.0-12.0	262			.0003	.0028	.0047	.0085	.0199	.0383	.0334	.0383	.36	*
12.0-13.0	249				.0027	.0042	.0085	.0176	.0277	.0311	.0277	.40	
13.0-14.0	231				.0027	.0054	.0085	.0140	.0256	.0275	.0256	.43	
14.0-15.0	207				.0020	.0042	.0073	.0142	.0242	.0265	.0242	.48	
15.0-16.0	187				.0018	.0040	.0071	.0126	.0194	.0247	.0194	.53	
16.0-17.0	170				.0019	.0042	.0077	.0162	.0192	.0289	.0192	.59	
17.0-18.0	156				.0021	.0049	.0092	.0172	.0234	.0314	.0234	.64	
18.0-19.0	138				.0026	.0045	.0069	.0142	.0150	.0261	.0150	.72	
19.0-20.0	115				.0017	.0035	.0071	.0150	.0219	.0271	.0219	.87	
20.0-21.0	100				.0015	.0046	.0099	.0157	.0198	.0306	.0198	1.00	
21.0-22.0	89				.0021	.0042	.0085	.0176	.0289	.0311	.0289	1.12	
22.0-23.0	75				.0014	.0035	.0064	.0163	.0269	.0277	.0269	1.33	
23.0-24.0	65				.0017	.0042	.0074	.0214	.0290	.0338	.0290	1.54	
24.0-25.0	46				.0019	.0033	.0058	.0222	.0255	.0330	.0255	2.17	
25.0-26.0	36				.0014	.0042	.0092	.0211	.0211	.0353	.0211	2.78	
26.0-27.0	19				.0011							5.26	
27.0-28.0	12				.0011							8.33	
28.0-29.0	8									.0240		12.50	*
29.0-30.0	6									.0031		16.67	*
30.0-31.0	2									.0008		50.00	*
31.0-32.0	2									.0046		50.00	*
32.0-33.0	2									.0027		50.00	*
33.0-34.0	2									.0073		50.00	*
34.0-35.0	2									.0008		50.00	*
35.0-36.0	2									.0157		50.00	*
36.0-37.0	1									.0074		50.00	*
37.0-38.0	1									.0008		100.0	*
38.0-39.0	1									.0153		100.0	*
39.0-40.0													
40.0-41.0													

review of the frequency tabulations showed that for the meridional shear the average empirical value amounts to 1.8 and for the zonal shear it is between 1.5 and 1.7. The factor 1.6 seems, therefore, a sound compromise between theory and practice. This takes care of departures from the normal distribution law and a mean value different from zero.

We have now established a theoretically acceptable critical value ( $\epsilon_c$ ) of 1.6 times  $2.28 \sigma$ , which equals  $3.65 \sigma$ . If the  $\sigma$  is known, we can easily compute this critical value  $\epsilon_c$ .

We also could use the 97.72 percentage value for the  $2.28 \sigma$  value. For the individual case, however, too many random variations may influence the result. Therefore, we may try to incorporate another procedure to decrease this effect. When we add the 84.1 percent and the 97.72 percent values, we obtain  $1.41\sigma$  plus  $2.28\sigma$ , which equals  $3.69\sigma$ . This is very close to  $3.65\sigma$ . Thus we may derive the critical value by employing the 84.1 and 97.72 percentage values.

Further consideration may be given to an observation tolerance error. An error of  $5 \text{ m. sec}^{-1}$  per 1000 m. for those extreme values seems to be within the limitation of measurements.<sup>3</sup> Thus we tolerate this error for the critical value and derive finally

<sup>3</sup> The first 3000 m. in table 2 are listed in 500-m. layers. It was decided for simplicity to adopt  $0.005 \text{ sec}^{-1}$  for those 500-m. layers, too.

$$\epsilon_c = P_{84.1} + P_{97.72} + 0.005 \text{ sec}^{-1}$$

where  $P_{84.1}$  is the 84.1 percentage value and  $P_{97.72}$  the 97.72 percentage value.

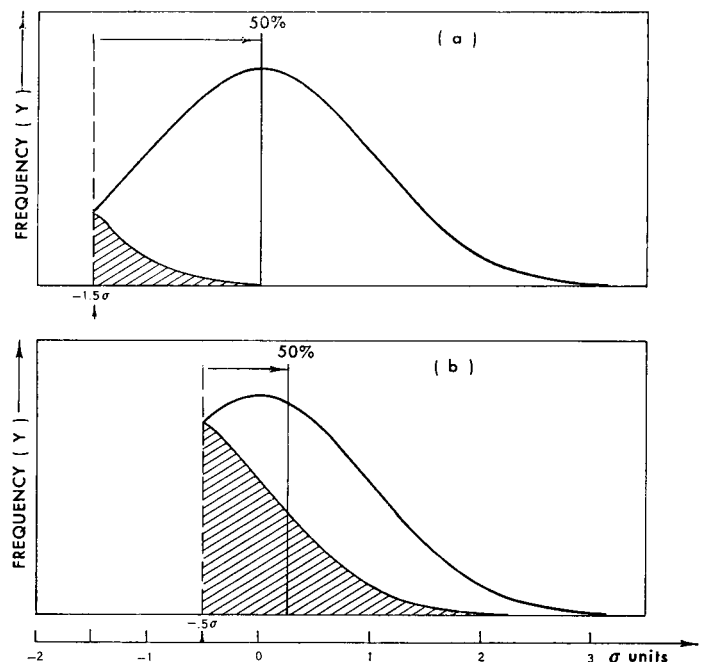


FIGURE 1.—Folding of normal distribution and relation to 50 percent line (median).

The unit of  $\epsilon_c$  is the same as in table 2, namely inverse seconds. From table 2 we may give a sample for computation of the critical value for the layer, surface to 0.5 km. It would be

$$\epsilon_c = .0202 + .0369 + .0050 = .0621 \text{ sec.}^{-1}$$

This value is listed in table 2 in the column titled "Essenwanger's sum". All shear observations exceeding this value as computed for each level were labeled suspicious. They are marked by an asterisk in table 2. Further description of the shear checking process will follow in section 4. Of the suspicious values, 85 percent had to be corrected, which is considered as high efficiency for this checking procedure.

Results of randomly picked shear values for checking have demonstrated that efficiency drops sharply for values below the critical value. Thus, by the outlined process we have achieved the goal to eliminate major errors, and tolerate (random) minor errors which have little bearing upon the determination of missile design criteria.

### 3. TEST OF DATA BY MAXIMUM WIND PROFILES

The highest and second highest (scalar) wind speeds for each of the stations at each of the 45 levels (or to the

highest level attained if it was less than 41 km.) were machine selected, and the observations<sup>4</sup> containing these high wind speed values were listed.

The verification of these speeds may be managed in two ways: verify or correct each and every value, or locate and correct the greater majority of the erroneous values, particularly values that are very large or appear to be inconsistent with a smooth profile which would be expected if all observations were correct. Practical economic considerations demanded that the latter be the guiding principle in the verification. One-third, or 8, of the stations received total verification; i.e., every high wind speed value was checked to provide a basis of comparison to determine the adequacy of the smooth profile verification procedure.

The verification procedure was divided into three steps or categories, namely: observation scan, terminating values, and profile scan. The sequence of the checking process might have been arbitrarily established, however, it appeared that the sequence listed above would provide maximum assurance that the final product, the maximum wind profile, was correct.

<sup>4</sup> The word "observation" in this report is used in the sense of characterizing the entire ascent as one taken observation and is in this way different from an observed value.

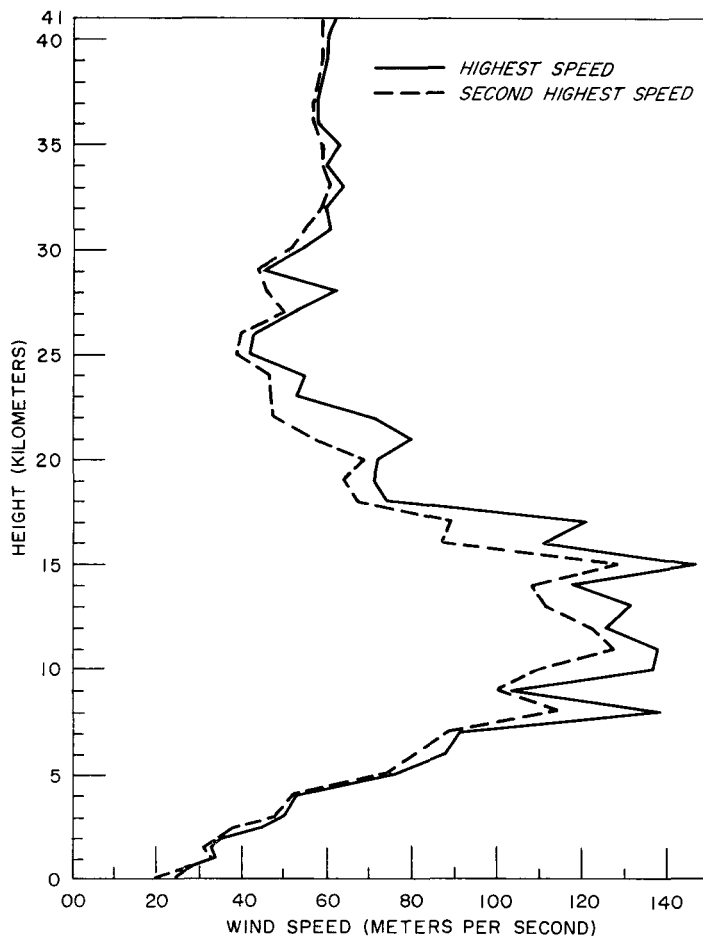


FIGURE 2.—Maximum wind profile before verification, station no. 1.

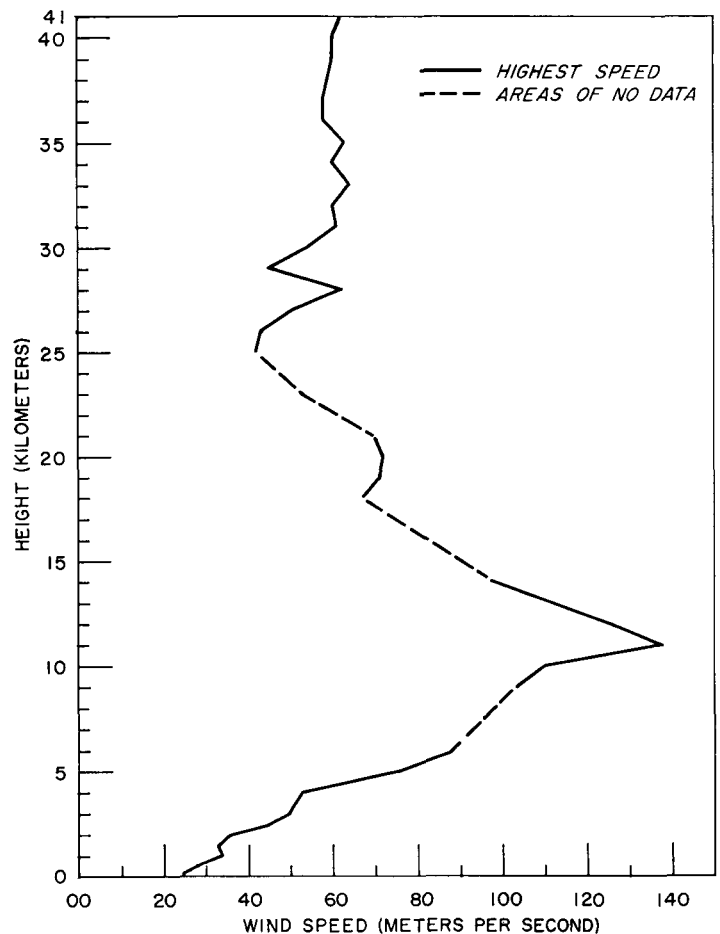


FIGURE 3.—Maximum wind profile after verification, station no. 1.

The observation scan was literally a visual scan of the observation. Each machine-selected observation was searched for apparent inconsistencies, such as a speed of 10 m.p.s. followed at the next level by a speed of 116 m.p.s., or similar rapid and large fluctuations in wind speed. Those observations containing such inconsistencies were checked and corrected as appropriate. This subjective scanning sufficed for errors detectable by discontinuity. Were all errors of this type, no further checking had been necessary. But though some profiles showed considerable smoothing from this process, most profiles still contained irregular contours which appeared suspicious. The checking process was therefore continued.

Each value occurring as the terminating speed for a given observation was considered suspicious. All observed wind values are the result of 2- or 4-minute averages except the last speed obtained which is allowed to be a 1-minute calculation if a 2- or 4-minute average is not available. Terminating values were therefore checked for representativeness. This eliminated the error from terminating fluctuations, but further smoothing was needed.

The profile scan necessitated the construction of the vertical profiles of these highest and second highest wind speeds. Figure 2 pictures the original profiles for station no. 1 to serve as an example. The first move was to apply to the profiles all corrections resulting from the first two

steps of the verification procedure. The profiles, after this preliminary correction, were then examined and questionable values were "picked off" for verification. It is obvious that a great deal of subjectivity was also encountered here, but a system was utilized. The values were chosen in sets, each set being verified and the resulting corrections applied before the next set was chosen. Each set consisted of those wind speed values which, if changed, would most smooth the contours of the profile. This set checking process was continued until no errors were found or so few were found that the profile was virtually unaffected.

Figure 3 shows the maximum wind profile after verification. Since this investigation dealt with only the highest and next to highest speeds, and values to replace these when they were deleted or changed to a value below that of the second highest were not included in this profile, thus dashed areas of no data appear. This meant the highest or second highest wind speed for a dashed level would have to be obtained by going back to the original records and selecting the now highest (or second highest) value. As this would have had to be done by hand at National Weather Records Center, it was decided to leave the profile as in figure 3. The missing values may easily be replaced by machine selection at Army Ballistic Missile Agency. The corrected profile, in

TABLE 3.—Maximum wind verification (observations checked and changed)

Station*	1	2	3	4	5	6	7	8	9	10	11	12
	Number of observations	Observations checked	Observations changed	Percent of observations checked Column 2 Column 1	Percent of observations changed Column 3 Column 1	Percent of observations checked, changed Column 3 Column 2	Number of profile points	Points checked	Points changed	Percent of points checked Column 8 Column 7	Percent of points changed Column 9 Column 7	Percent of points checked, changed Column 9 Column 8
By procedure:												
1.....	62	31	23	50	37	74	90	40	23	44	26	58
2.....	68	38	27	56	40	71	90	48	29	53	32	60
3.....	69	30	3	43	4	10	64	40	3	63	5	8
4.....	58	25	18	43	31	72	90	55	23	61	26	42
5.....	75	24	9	32	12	38	72	40	17	56	24	43
6.....	48	27	19	56	40	70	90	69	38	77	41	55
7.....	51	21	9	41	18	43	90	53	11	59	12	21
8.....	56	40	23	71	41	58	90	69	37	77	41	54
9.....	48	27	6	56	13	22	86	56	6	65	7	11
10.....	86	38	8	44	9	21	64	37	8	58	13	22
11.....	64	36	18	56	28	50	90	61	24	68	27	39
12.....	50	21	9	42	18	43	65	36	10	55	15	28
13.....	60	32	19	53	32	59	90	58	32	64	36	55
14.....	86	35	20	41	23	57	75	39	22	52	30	56
15.....	73	45	31	62	42	69	90	63	40	70	45	63
16.....	72	42	33	58	46	79	90	67	49	74	54	73
Totals.....	1,026	512	275	50	27	54	1,326	831	372	63	28	45
By total verification:												
17.....	44	44	23	100	52	52	90	90	36	100	40	40
18.....	64	64	37	100	58	58	70	70	35	100	50	50
19.....	60	60	45	100	75	75	90	90	60	100	67	67
20.....	73	73	46	100	63	63	76	76	40	100	53	53
21.....	55	55	42	100	76	76	86	86	59	100	69	69
22.....	58	58	28	100	48	48	73	73	39	100	53	53
23.....	50	50	30	100	60	60	73	73	43	100	59	59
24.....	59	59	32	100	54	54	72	72	26	100	36	36
Totals.....	463	463	283	100	61	61	630	630	338	100	54	54
Grand Totals.....	1,489	975	558	65	37	57	1,956	1,461	710	75	36	49

\*The sequence of stations is not identical with station listing in table 1.

general, takes on a smoother appearance with one distinct layer of maximum wind speed. This is true with all of the stations.

Table 3 denotes the amount of checking and changing done in the maximum wind profile verification program. The data are subdivided into two groups; the first being for the 16 stations for which the verification was accomplished by the procedure described in the preceding paragraphs, the second for the 8 stations which received total verification. The combination of these yields the overall results. Table 3 is self explanatory.

As can be seen from table 3, there is some question concerning the efficiency of this three-step procedure, particularly when the percentage of values changed becomes very large. It would seem that there still are a significant number of erroneous observations unchecked. It must be remembered, however, that the criterion for terminating verification was that the contours of the profile remain virtually unchanged, since the original goal of the verification was to obtain a representative maximum wind profile. This goal was attained in all cases where the profile was not completely destroyed. Three stations lacked sufficient data and for one the entire profile would have needed reconstruction.

This method was designed for adequacy, with efficiency being second in importance. Methods aspiring toward high efficiency may be similar to the above procedures but, would necessarily include more objectivity and less subjectivity. The method outlined for vertical wind shear verification with objective selection may be adaptable to wind speeds using similar frequency distributions.

#### 4. TEST OF DATA BY WIND SHEAR

Selections were made from the shear distribution tables similar to table 2. These tables list the vertical shear by 1-km. layers, beginning at 3 km. Below 3 km., 500-m. layers were taken. The first attempt at selection was by picking out values that seemed to be erratic. This was abandoned almost immediately since it was too laborious and uncertain (nearly 14,000 observations contained all of the maximum shear values).

Then it was decided to use the method outlined in section 2. This method is as follows. Obtain the sum of the shear values at the cumulative percentage frequencies of 84.1 percent and 97.72 percent plus  $0.0050 \text{ sec.}^{-1}$ , then select shears for checking on the basis that any shear value greater than this sum is suspicious. In addition, a selection by observations was made: any value that had been obtained at a level reached ten or less times during the period of record was considered as suspicious, regardless of magnitude.

The selection of the suspicious values is demonstrated in table 2. This tabular form portrays layer versus cumulative percentage frequency with supplementary columns for observations counts and the observed maximum shear. The vertical wind shear (henceforth shear) values were

computed for both the zonal and meridional components. Listings of the profile for the maximum wind shear accompanied the listings of the form of table 2.

In the first set of tabulations maximum shear values only were subjected to verification. Later all shear values exceeding the "critical" threshold (Essenwanger's sum) were listed as a review program. It was understood that, regardless of what portion or portions of an observation first raised suspicion, the entire observation was to be subjected to a checking process.

Secondary review was limited to a review check program which was designed to encompass not only highly suspicious observations not contained in the first phase, but also observations in which errors may have occurred or been overlooked during previous verification. The latter, fortunately, occurred very few times. Even with this, some few errors undoubtedly still escaped detection.

The combination of these methods has proven acceptably efficient in that 85 percent of the observations selected as suspicious were found to be erroneous.

The errors have been tabulated in three categories: clerical, instrumental, and computational. The clerical errors were based on the premise that no technical observer training was required to perform the work classified as clerical. They were subdivided into three types: punching, extraction (or transcription), and plotting of Form WBAN 20A. The punching was, of course, the production of the cards constituting the original card decks. The extraction was the "picking off" of values from the WBAN 20A, a plot of the observed data in the form of wind speed and direction versus height. Plotting refers to wrong plots of WBAN 20A.

Instrumental error may be defined as large and rapid fluctuations in angles (azimuth and elevation) incompatible with the calculated height changes.

The computational error was subdivided into two types: Calculations on WBAN 20 (the observer's work sheet) and fictitious ascension rates (of the balloon) as calculated from erroneous pressure-temperature-time measurements. Fictitious ascension rates may be thought of as instrumental, but only the tracking equipment was considered an instrument in this study. The tracking equipment consisted mainly of the theodolite (visual tracking), the SCR-658 (manual radiosonde tracking), and the GMD-1 and GMD-1A (automatic radiosonde tracking).

Table 4 shows the error statistics resulting from the verification of the shear selected observations. The headings are self explanatory. Some observations contained more than one type of error so that the occurrence of errors exceeds the number of observations changed. The end results in this tabulation prove interesting in the predominance of the instrumental and clerical errors. The relatively small computational error is gratifying.

We note in table 4 that from all observations only 3.5 percent were found to contain suspicious values; from the suspicious observations 85 percent, i.e., 2.9 percent of all

TABLE 4.—Vertical wind shear verification (error statistics)

Station*	Observations				Number of errors				Percentage of				Percentage of total errors		
	1	2	3	4	5	6	7	8	Total observations containing maximum shear	Total observations containing suspicious	Suspicious changed	Changed of total	Clerical	Instrumental	Computation
	Total number	Containing maximum shear	Suspicious	Changed	Total	Clerical	Instrumental	Computation	Column 2 Column 1	Column 3 Column 1	Column 4 Column 3	Column 4 Column 1	Column 6 Column 5	Column 7 Column 5	Column 8 Column 5
1.....	5,114	595	226	159	167	92	57	18	11.6	4.4	70	3.1	55	34	11
2.....	4,992	556	223	194	249	101	93	55	11.1	4.5	87	3.9	41	37	22
3.....	5,114	537	169	133	134	27	100	7	10.5	3.3	80	2.6	20	75	5
4.....	5,114	626	208	159	166	78	58	30	12.2	4.1	76	3.1	47	35	18
5.....	3,652	550	106	87	107	33	50	24	15.1	2.9	82	2.4	31	47	22
6.....	5,114	602	215	195	219	106	73	40	11.8	4.2	91	3.8	48	34	18
7.....	1,462	515	112	83	97	38	15	44	35.2	7.7	74	5.7	39	16	45
8.....	5,054	542	146	118	152	47	53	52	10.7	2.9	81	2.3	31	35	34
9.....	3,896	504	124	91	102	72	26	4	12.9	3.2	73	2.3	71	25	4
10.....	5,114	608	103	80	75	37	35	3	11.9	2.0	78	1.6	49	47	4
11.....	5,114	627	178	162	189	60	66	63	12.3	3.5	91	3.2	32	35	33
12.....	5,114	535	182	121	152	77	58	17	10.5	3.6	66	2.4	51	38	11
13.....	5,114	612	182	167	212	84	69	59	12.0	3.6	92	3.3	40	32	28
14.....	4,686	619	112	90	121	39	48	34	13.2	2.4	80	1.9	32	40	28
15.....	4,990	622	150	132	167	73	60	34	12.5	3.0	88	2.6	44	36	20
16.....	5,114	640	226	219	287	132	86	69	12.5	4.4	97	4.3	46	30	24
17.....	5,114	650	223	195	267	105	104	58	12.7	4.4	87	3.8	39	39	22
18.....	5,114	591	129	117	127	24	91	12	11.6	2.5	91	2.3	19	72	9
19.....	5,114	541	251	237	300	129	120	51	10.6	4.9	94	4.6	43	40	17
20.....	5,114	549	127	116	121	9	104	8	10.7	2.5	91	2.3	7	86	7
21.....	5,114	486	217	198	227	47	139	41	9.5	4.2	91	3.9	21	61	18
22.....	5,114	509	77	61	64	12	46	6	10.0	1.5	79	1.2	19	72	9
23.....	5,114	521	162	146	164	39	113	12	10.2	3.2	90	2.8	24	69	7
24.....	3,652	502	103	80	93	35	49	9	13.7	2.8	78	2.2	38	52	10
Grand total.....	114,208	13,639	3,951	3,340	3,959	1,496	1,713	750	11.9	3.5	85	2.9	38	43	19

\*The sequence of stations is not identical with station listing in table 1.

observations, had to be changed. This means that one or several values had to be corrected in 2.9 percent of the observations. As one observation contains numerous levels, the actual percentage of errors in relation to the wind data of all levels is far less. The clerical error contributed 1.1 percent, the instrumental error 1.3 percent, and the computational error 0.5 percent. A perfect card deck would be the ideal goal, of course. As one should expect, this goal cannot be reached without a thorough check of the produced data. Thus, the magnitude of percentage of the larger error discovered by the method outlined above stays well within reasonable tolerance limits.

A few brief remarks may be in order concerning the difficulties involved in verification. First, the major division of errors is not as clear-cut as it appears. Prior to 1956, the extracted values were not always entered in the allotted space on WBAN 20A. Errors occurring subsequent to plotting were attributed to extraction in such cases. Then, beginning on January 1, 1956, WBAN 20A was no longer a filed form so that plotting errors were undetectable. WBAN 20B had come into being, on which was allotted space for the "extracted" values. Again, arbitrarily, errors occurring between calculation (WBAN 20) and punching were attributed to extraction, so that an overlap occurred. Fortunately, plotting errors contributed very little to the mass of clerical errors.

Perhaps the most difficult problems arose with erratic instrument data and fictitious ascension rates. Balloon heights are determined through pressure-temperature

considerations on Forms WBAN 31 A and B. Erroneously low pressure or high temperature readings result in erroneously high height determinations, which when considered with ascension time, may present unduly large ascension rates. This normally occurred in the topmost layers of the higher ascents. This same reasoning can also account for highly erratic or incompatible height changes at any level with the erroneous pressure or temperatures being high or low. Highly erratic angles are caused mainly by equipment malfunction and limitation. Any appreciable influence of turbulence would be confined to the lowest levels except under rare conditions such as balloon entry into a thunderstorm or possibly clear air turbulence. The first event was eliminated by consideration of the weather reports. The latter event, generally not too frequent in occurrence, is not known at the present in sufficient detail to make an unequivocal decision. It was felt, however, that when low ascension rates are involved, it is more likely that the data are erroneous due to instrumental errors. This conclusion may be due for revision after knowledge of clear air turbulence has improved.

In general, erratic data were determined subjectively since there were no adequate objective methods available. Ascension rates were, in general, declared fictitious in a subjective manner since time limitations and unavailable observation data precluded thorough checking. There are two compensating factors, however, in that observation data were always given the benefit of the doubt, and all decisions were made by qualified meteorologists.

### 5. CONCLUSION

This study has discussed the possibility of checking wind data by maximum wind profiles and wind shear distributions. While the maximum wind profiles were evaluated for suspicious values by profile scan, the checking process by wind shears was based upon computation of a critical value  $\epsilon_c$ . Exceedance of this critical value made the shear value suspicious and subject to verification.

The derivation of the critical value  $\epsilon_c$  was developed and the application to 24 stations showed an efficiency of 85 percent, which may be considered very high. Although 3.5 percent of the observations proved to be suspicious, and 2.9 percent had to be corrected, the actual corrections are less, as one observation in the average contains between 20 and 30 level values, not all of which had to be corrected.

The errors were traced and divided into 1.1 percent clerical errors, 1.3 percent instrumental errors, and 0.5 percent computational errors. These are within reasonable limitation.

It may be stressed that establishment of maximum wind profiles or maximum wind shears may be better approached by theoretical statistical processes in order to

eliminate the effect of the relatively short period of available data record. Time and cost limitations, however, prevented further investigation in this direction.

### ACKNOWLEDGMENTS

The authors wish to express their special appreciation to Mr. Earl M. Ritchie of NWRC, who started work on the maximum wind speed procedure and completed most of this phase of the project. We would like also to recognize the outstanding efforts of Mr. Norman Graham of NWRC for the technical work in the accomplishment of the project.

### REFERENCES

1. E. J. Gumbel and P. G. Carlson, "Extreme Values in Aeronautics," *Journal of the Aeronautical Sciences*, vol. 21, No. 6, June 1954, pp. 389-398.
2. H. C. S. Thom, "Frequency of Maximum Wind Speeds," *Proceedings of the American Society of Civil Engineers*, Sgs. No. 539, 1954, pp. 1-11.
3. H. C. S. Thom, "Distribution of Extreme Winds in the United States," *Journal of the Structural Division, Proceedings of the American Society of Civil Engineers*, vol. 86, No. ST 4, Apr. 1960, pp. 11-24.